ASTRONOMICAL DATA ANALYSIS SOFTWARE AND SYSTEMS XIV
ASP Conference Series, Vol. 347, 2005
P. L. Shopbell, M. C. Britton, and R. Ebert, eds.

Linearization of Spitzer IRS Data Via Minimization of χ^2 With Correlated Errors

John W. Fowler

Spitzer Science Center, MS 314-6
California Institute of Technology, Pasadena, CA 91125

Abstract. The Spitzer Infrared Spectrograph (IRS) data are taken via read-without-reset measurements to obtain multiple samples forming a photometric "ramp" for each pixel in an echellogram. Each ramp is linearized via a quadratic model. After linearization, a quality-assurance test is performed to determine how linear each pixel's ramp has become. This is accomplished by fitting a straight line to the ramp via χ^2 minimization. The goodness of fit is of primary importance, since this determines whether the inevitable deviations from linearity are statistically significant given the estimated photometric noise. Because the latter is dominated by photon noise which is summed up the ramp, the χ^2 parameter used to measure goodness of fit must include the effects of correlated errors. This paper describes the construction of the full error covariance matrix and its use in the χ^2 minimization.

1. Introduction

One of the challenges in the scientific application of new hardware technology is quantifying the extent to which the hardware behavior is understood. The depth of this understanding is invariably revealed most accurately by the state of the corresponding data-analysis algorithms devoted to the "error analysis", i.e., the mathematical modeling of the uncertainty in the values of any parameters derived from measurements obtained via the hardware. This uncertainty may originate in both the object of measurement (e.g., fluctuations in the arrival rate of photons) and the hardware itself (e.g., dark current drifts). The characterization of uncertainty determines the scientific usefulness of the measurements, so it is desirable to model the hardware as accurately as possible in order to minimize the uncertainty, then to evaluate the irreducible uncertainty accurately enough to make reliable statistical interpretations (e.g., to be confident in judging the difference between an expectable fluctuation and an externally generated "glitch" such as a cosmic ray hit).

It may happen that the errors in two measurements are correlated through mutual dependence on common random events. This can significantly affect statistical interpretations and thus must be taken into account in the error model. A typical example arises in the evaluation of "goodness-of-fit" parameters, the most common of which is the χ^2 statistic. In practice, the effect of ignoring correlations when computing χ^2 is usually underestimation, which tends to suggest misleadingly that the errors are slightly overestimated and thus comfortably conservative. Including the correlations could reveal that the statistical significance of the discrepancies is actually larger than expected, indicating possible

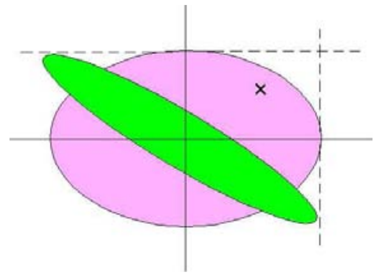

Figure 1. The error ellipse for correlated errors is rotated and has a smaller area than the error ellipse whose covariance matrix has the same diagonal elements but zero off-diagonal elements. Ignoring correlation would underestimate the significance of the deviation of the point marked by X.

problems with the modeling. In any case, an accurate χ^2 value is highly desirable for its ability to signal most problems with the hardware model or the error estimation.

An easily visualized example of correlated errors is that of celestial coordinates whose uncertainty involves an error ellipse that is rotated with respect to the coordinate system. Figure 1 shows a case in which a positive error on the horizontal axis is more likely to accompany a negative error on the vertical axis, i.e., the errors are negatively correlated. The smaller error ellipse includes the correlation information, while the larger unrotated ellipse corresponds to an error covariance matrix with the same diagonal elements but zero off-diagonal elements. The area inside a one-sigma contour is clearly greater when the off-diagonal elements are taken to be zero.

2. Correlated Errors in Spitzer Infrared Spectrograph Ramps

The Spitzer Infrared Spectrograph employs several echelle spectrographs whose measurements are formatted as FITS data cubes, i.e., a stack of N 128×128 image planes, where $4 \leq N \leq 32$. Each pixel has N nondestructive readouts in the data cube. Since photo-electrons accumulate between each readout, the pixel values generally increase with plane number, and so the set of values is called a "ramp". Ideally each ramp would take the form of a straight line, but noise and nonlinear response prevent perfect linearity. The nonlinear response can be well approximated as quadratic, and each pixel is calibrated accordingly. This calibration is applied as part of the data reduction process, i.e., the quadratic model is inverted to "linearize" the ramp. For quality assessment purposes, the linearized ramp is fit to a straight line, and three goodness-of-fit parameters are evaluated: the linear correlation coefficient, RMS dispersion, and χ^2 statistic.

χ^2 depends on the deviations of the linearized data from the best-fit straight line and on the *a priori* uncertainties in these data. The uncertainties are provided by an error model that incorporates such instrumental effects as calibrated read noise, dark subtraction uncertainty, linearization uncertainty, etc., and "photon noise" caused by "counting-statistics" fluctuations in the incoming photon stream. The photon noise is well approximated as a Poisson process, and the sum of the other uncertainties is well approximated as a Gaussian process.

The Poisson process is generally well into its Gaussian limit also, but the Poisson character is retained in that the variance is equal to the mean, where the mean is the measured number of photo-electrons in the detector well at readout time. Errors in this mean due to Poisson fluctuations at one readout are inherited at the next readout, and so each readout above the first contains the sum of all fluctuation errors below, correlating these errors. The other sources of error occur downstream from the well and are effectively independent from one readout to the next. The error in readout no. i of a ramp is

$$\varepsilon_i = \varepsilon_{pi} + \varepsilon_{ui} = \sum_{j=1}^{i} \Delta\varepsilon_{pj} + \varepsilon_{ui}$$

where ε_{pi} is the total Poisson photon noise at plane i, i.e., the sum of the incremental fluctuations $\Delta\varepsilon_{pj}$ over all lower planes up to and including plane i, and ε_{ui} is the total uncorrelated noise at plane i. We are interested in the expectation value of the product of the errors at any two planes m and n:

$$\varepsilon_m \varepsilon_n = \left(\sum_{j=1}^{m} \Delta\varepsilon_{pj} + \varepsilon_{um} \right) \left(\sum_{k=1}^{n} \Delta\varepsilon_{pk} + \varepsilon_{un} \right)$$

$$= \sum_{j=1}^{m} \left(\Delta\varepsilon_{pj} \sum_{k=1}^{n} \Delta\varepsilon_{pk} \right) + \varepsilon_{um}\varepsilon_{un} + \varepsilon_{um} \sum_{k=1}^{n} \Delta\varepsilon_{pk} + \varepsilon_{un} \sum_{j=1}^{m} \Delta\varepsilon_{pj}$$

Since ε_{um} is uncorrelated with all errors in the ramp other than itself (i.e., ε_{un} when $n = m$), the last two terms on the right will become zero when we take expectation values. Similarly, the second term's expectation value will be zero for $m \neq n$. Furthermore, each incremental photon-noise error $\Delta\varepsilon_{pj}$ is uncorrelated with each other $\Delta\varepsilon_{pk}$ except for $j = k$. For $m = n$, the expectation values are therefore

$$\left\langle \varepsilon_n^2 \right\rangle \equiv v_n = \left\langle \sum_{k=1}^{n} \Delta\varepsilon_{pk}^2 \right\rangle + \left\langle \varepsilon_{un}^2 \right\rangle = \sigma_{pn}^2 + \sigma_{un}^2$$

where v_n is the total error variance at plane n (i.e., v_{nn} with the second index suppressed), the sum of the photon noise variance σ_{pn}^2 and the total uncorrelated noise variance σ_{un}^2 at that plane. This total error variance is computed from the error model and is available at each processing stage. What is not provided and must be reconstructed are the off-diagonal elements of the error covariance matrix, v_{mn}, which we obtain by considering $m > n$:

$$\langle \varepsilon_m \varepsilon_n \rangle \equiv v_{mn} = \left\langle \sum_{k=1}^{n} \Delta\varepsilon_{pk}^2 \right\rangle = \sigma_{pn}^2$$

So the covariance of the error at plane m with that at the lower plane n is the photon noise at plane n. Because of the Poisson character of the photon noise, this is just the number of electrons at plane n, which we will denote y_n to be consistent with the linear equation used in the fitting,

$$y = ax + b$$

where x is the plane number.

3. Linear Fit via χ^2 Minimization With the Full Error Covariance Matrix

We can now construct the full error covariance matrix. The diagonal elements v_n are supplied by the error model, and the off-diagonal elements v_{mn} are just y_n, where $m > n$. For example, a five-plane ramp would have the error covariance matrix:

$$\Omega = \begin{pmatrix} v_1 & y_1 & y_1 & y_1 & y_1 \\ y_1 & v_2 & y_2 & y_2 & y_2 \\ y_1 & y_2 & v_3 & y_3 & y_3 \\ y_1 & y_2 & y_3 & v_4 & y_4 \\ y_1 & y_2 & y_3 & y_4 & v_5 \end{pmatrix}$$

We define the vector u whose i^{th} component is the fitting deviation at plane i:

$$u_i \equiv y_i - ax_i - b$$

With a full $N \times N$ error covariance matrix, χ^2 is

$$\chi^2 = uWu^T$$

where u^T is the transpose of u, and W is the inverse of Ω and has elements w_{ij}. Expanding the vector-matrix-vector multiplication yields

$$\chi^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}u_i u_j$$

Inserting the definition of the components of u, differentiating with respect to the fitting coefficients a and b, and setting the results to zero gives a 2×2 system of linear equations

$$a\sum_{i=1}^{N} w_i x_i + b\sum_{i=1}^{N} w_i = \sum_{i=1}^{N} w_i y_i$$

$$a\sum_{i=1}^{N} x_i z_i + b\sum_{i=1}^{N} w_i x_i = \sum_{i=1}^{N} y_i z_i$$

where

$$w_i \equiv \sum_{j=1}^{N} w_{ij},\, z_i \equiv \sum_{j=1}^{N} w_{ij} x_j$$

This system of equations is easily solved for a and b, and then the value of χ^2 is computed from the equation defining it above.

When applied to well behaved simulation data, the calculations described above reveal the typical features of curve fitting via χ^2 minimization with the inclusion of correlated errors: the coefficients obtained are rather insensitive to whether the error correlation is taken into account, but the value of χ^2 tends to be highly dependent on it. As Figure 1 shows, while the significance of a deviation depends noticeably on which error ellipse is used in the interpretation, the weighted average of many points spread randomly according to the smaller ellipse would not tend to depend strongly on which ellipse supplied the inverse-variance weighting.

Simulation data were generated for five 16-plane ramps with illumination designed to sample the dynamic range of real IRS pixels. The nominal increments in electrons per plane varied linearly from 2,250 to 20,250 from pixel 1 to pixel 5. With a gain of 5.0, this produces a DN of 64,800 at plane 16 of the brightest pixel, close to the limit of an unsigned 16-bit integer. A nominal read noise of 10 electrons/read supplied the uncorrelated non-photon noise. Pseudorandom noises were added to the nominal signal to generate the ramps, and then the linear fits were computed. Numerous trials were performed, and the results shown in the table below are typical.

Table 1. Linear Fits With and Without Error Correlation Included

Pixel	$\Delta e^-/Plane$	χ^2_{uncorr}/N_{df}	$P(\chi^2_{uncorr})$	χ^2_{corr}/N_{df}	$P(\chi^2_{corr})$
1	2250	0.31	0.00719	1.41	0.86001
2	6750	0.24	0.00173	0.86	0.39928
3	11250	0.26	0.00294	1.32	0.81334
4	15750	0.42	0.03064	1.09	0.63516
5	20250	0.29	0.00474	0.98	0.52799

Notes: N_{df} is the number of degrees of freedom, here 14, since there are 16 planes and 2 fitting coefficients; the expectation value of χ^2/N_{df} is unity, and the significance of a given deviation from unity increases with N_{df}; $P(\chi^2)$ is the fraction of all χ^2 that have N_{df} degrees of freedom and are less than or equal to the argument.

For all pixels, χ^2 is too small when error correlation is ignored. $P(\chi^2)$ should be approximately uniformly distributed over multiple trials, and this is clearly not the case unless the error correlation is taken into account. Similar trends are seen in real flight data, but so far these are somewhat diluted by the presence of many uncorrelated errors, some of which are known to be overestimated. Work is underway to refine the error models.

Acknowledgments. This work was performed at the Spitzer Science Center as part of a mission/project managed by Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.